Price Dispersion in OTC Markets: A New Measure of Liquidity

First draft: February 2008
This draft: April 2008

Abstract

   In this paper, we model price dispersion effects in over-the-counter (OTC) markets to show that in the presence of inventory risk for dealers and search costs for investors, traded prices may deviate from the expected market valuation of an asset. We interpret this deviation as a liquidity effect and develop a new liquidity measure quantifying the price dispersion in the context of the US corporate bond market. This market offers a unique opportunity to study liquidity effects since, from October 2004 onwards, all OTC transactions in this market have to be reported to a common database known as the Trade Reporting and Compliance Engine (TRACE). Furthermore, market-wide average price quotes are available from Markit Group Limited, a financial information provider. Thus, it is possible, for the first time, to directly observe deviations between transaction prices and the expected market valuation of securities. We quantify and analyze our new liquidity measure for this market and find significant price dispersion effects that cannot be simply captured by bid-ask spreads. We show that our new measure is indeed related to liquidity by regressing it on commonly-used liquidity proxies and find a strong relation between our proposed liquidity measure and bond characteristics, as well as trading activity variables. Furthermore, we evaluate the reliability of end-of-day marks that traders use to value their positions. Our evidence suggests that the price deviations are significantly larger and more volatile than previously assumed. Overall, the results presented here improve our understanding of the drivers of liquidity and are important for many applications in OTC markets, in general.

Keywords: liquidity, corporate bonds, market microstructure, OTC markets.

JEL classification: G12.

Price Dispersion in OTC Markets: A New Measure of Liquidity

1   Introduction

The liquidity of financial markets is of crucial importance for diverse market participants such as corporations, investors, broker-dealers, as well as regulators. While there is an extensive literature on liquidity effects in exchange-traded markets, particulary those for equities, there is very little research, thus far, on these effects in *over-the-counter* (OTC) markets. Liquidity and its effects on prices have to be considered in all investment decisions, and this issue seems to be of special importance for *illiquid markets*, particularly OTC markets, where prices are the result of bilateral negotiations between investors and dealers. The objective of this paper is to bridge this gap by developing a tractable, theoretical model for OTC markets and testing its implications empirically with US corporate bond market data. This market offers a unique opportunity to study liquidity effects since, from October 2004 onwards, all OTC transactions have to be reported to a centralized database known as the Trade Reporting and Compliance Engine (TRACE). In addition, there is a valuation service provided by Markit Group Limited, which surveys broker-dealers at the end of each trading day to obtain a composite quote for each security. The combination of these two data-sets offers us an opportunity to construct a metric of price dispersion in an OTC market, and to compare these results to trading measures such as bid-ask spreads, which are often employed as proxies for this information.

OTC markets are especially interesting from the perspective of liquidity because of their trading architecture. In the absence of a centralized trading platform, buy and sell transactions have to be directly negotiated by agents who need to contact one of the potential dealers in the market. Although bid-ask quotations are normally posted by dealers, e.g. on Bloomberg or Reuters, these are not binding, i.e., they often only hold for small quantities or can be stale in some cases. Thus, investors potentially have to negotiate with multiple dealers to trade at an acceptable price. This market structure is very different from exchange-traded markets where a central order book is available to all market participants. Even so, one might expect that in the absence of any market frictions, traded prices would still be equal to the expected market valuation in OTC markets. Obviously, this is not the case in the presence of market frictions, such as fixed costs and inventory risk for dealers, and search costs for investors. These market imperfections could lead to traded prices that are potentially higher or lower than the market valuation of a particular instrument. They could even result in situations where the instrument is traded at significantly different prices, at approximately the same time. Therefore, these price dispersions are of interest

when analyzing market liquidity. However, market-wide transaction data are generally not available for OTC markets. A rare exception is the US corporate bond market, where the transaction data are collected in a centralized database. Thus, our paper is especially interesting, since price dispersion effects between transaction prices and aggregated market valuations of securities are *directly* observable in the US corporate bond market, using the TRACE and Markit data.

Our paper makes three contributions to the literature on market microstructure and liquidity in the context of OTC markets. First, we develop a new measure of liquidity based on price dispersion effects, which we derive from our model. This measure is the root mean squared difference between the traded prices of a particular bond, provided by TRACE, and its respective market valuation, provided by Markit. Thus, it is an estimate of the absolute deviation, and, more importantly, can be interpreted as the volatility of the price dispersion. An analysis at the level of the aggregate market, as well as at the bond level, shows that this price dispersion is significantly larger than quoted bid-ask spreads and also shows more variation across bonds. This indicates that the overall liquidity of the corporate bond market is rather low, and that liquidity, so far, could only be roughly approximated by quotations.

Second, we relate our price dispersion measure to conventional liquidity proxies. Thus, we show that the new measure is indeed related to liquidity and identify reliable proxies, which are especially important for OTC markets, where traded prices are not easily available. For this purpose, we test whether our measure can be related to differential liquidity at the bond level, by relating it to commonly-used liquidity proxies, i.e., bond characteristics, trading activity variables, as well as established liquidity measures used in the academic literature. The resulting regression models show high explanatory power, and the effects remain stable over time, indicating a strong relation between the price dispersion and liquidity-related variables. According to our results, the most important liquidity proxies for corporate bonds are the amount issued, maturity, age, rating, bid-ask spread, and trading volume, as well as the price impact measure introduced by Amihud (2002). In the event, we find a strong relation between our liquidity measure, and bond characteristics and trading activity variables. Therefore, our measure can potentially be used to extract the liquidity component of corporate bond yield spreads.

Third, our results serve as an indirect test of the reliability of end-of-day marks provided by average prices or bid-ask quotations. Using a volume-weighted hit-rate analysis, we find that only 51.12% of the TRACE prices and 58.59% of the Markit quotations lie within the bid and ask range quoted on Bloomberg. These numbers are far smaller than previously assumed. Since these marks are widely used in the financial services industry, our findings may be of interest to financial institutions and their regulators.

In the theoretical section of the paper, we develop a market microstructure model focusing on the price dispersion effects in OTC markets. As argued earlier, these deviations can be interpreted as the effect of liquidity in the presence of inventory and search costs. In this setting, a perfectly liquid market is characterized by negligible deviations of traded prices from their market valuations, whereas illiquid markets show large dispersions. Investors will perceive these deviations as indicative of the transaction costs of trading, and consider them when making investment decisions. In the market microstructure literature, price dispersion effects are explained by either fixed costs and inventory risk for dealers and by search costs for investors, or as arising due to asymmetry of information across traders and dealers.[2]

Garbade and Silber (1976) present one of the earliest price dispersion models in the context of the US Treasury bond market. In their setup, investors with search costs are confronted with an exogenously given probability distribution of potentially offered prices, when contacting an arbitrary dealer. Consequently, investors will accept deviations from the perceived fundamental value, up to a certain point, to avoid the marginal search costs arising from contacting an additional dealer. Similar ideas are put forward in Garman (1976) and in Amihud and Mendelson (1980), where a centralized dealer with inventory risk is confronted with stochastic arrivals of investors' offers. The optimal inventory and price setting policy are derived from the tradeoff between risk and return for each agent. Ho and Stoll (1980) and Ho and Stoll (1983) focus on the competition among dealers deriving equilibrium inventories and market spreads. Grossman and Miller (1988) model liquidity events for risk-averse investors resulting in an immediate need to trade the security. The investor can trade immediately by incurring a cost, or wait one period, bearing the risk of adverse price movements. This tradeoff directly yields a liquidity cost for immediacy and the optimal number of dealers. Bagehot (1971), Glosten and Milgrom (1985), Kyle (1985), and Easley and O'Hara (1987), as well as many others who followed, introduce the concept of informed traders versus liquidity traders and interpret bid-ask spreads as the compensation for adverse selection. Huang and Stoll (1997) provide a simple model combining these different effects, and show, based on stock market data, that order processing costs and inventory risk are the important components of transaction costs. In a related paper, Hansch et al. (1998) show with data from the London Stock Exchange that inventory determines dealer behavior. Bollen et al. (2004) provide a model which includes other microstructure effects, such as the minimum tick size, the inverse of trading volume, competition among dealers, and expected inventory holding premium. They demonstrate that their model performs well for Nasdaq data.

Most of the aforementioned literature is in the context of a framework with a market-maker (or multiple

---

[2]See Amihud et al. (2006) for a comprehensive survey of the literature on liquidity.

market-makers) with a centralized order book, an abstraction for an exchange-traded market. Turning to models for OTC markets, Duffie et al. (2005) and Duffie et al. (2007) present a market with risk-neutral investors who face stochastic holding costs, generating trading necessity. The availability of dealers and investors is modeled by their respective trading intensities. The search time and relative bargaining power determines equilibrium prices in their model.

Following this line of the literature, we develop a tractable, but sufficiently realistic model, to capture liquidity effects in an OTC market. To this end, we explicitly model the stochastic inventories of multiple dealers and their capital costs/constraints, together with the search costs of investors, abstracting from issues relating to information asymmetry and adverse selection. This simple, yet realistic, formulation allows us to obtain a clear interpretation for information observable by investors in OTC markets in our symmetric information context. Our model relates these information sources to each other and enables us to measure the degree of liquidity using TRACE prices and Markit quotations.

Building on this setup, in the empirical implementation of our model, we quantify our measure of price dispersion in the context of the US corporate bond market, one of the largest OTC markets in the world. Our data set covers 1,800 bonds with 3,889,017 observed transaction prices for the time period October 1, 2004 to October 31, 2006. This market is especially interesting for our purposes, as liquidity differences across individual bonds seem to be rather pronounced: very few bonds are traded frequently, while most other bonds are almost never traded at all.[3] Differences in inventory risk and search costs are, therefore, evident in this market, making it an ideal laboratory to test our model. We use our data set to analyze the price dispersion effect and its relation to liquidity at the level of the aggregate market and individual bonds.

The empirical literature suggests a whole range of liquidity proxies in the context of corporate bond markets. Several authors study the impact of liquidity, based on corporate yields or yield spreads over a riskfree benchmark. Most of these papers rely on indirect proxies such as the coupon, age, amount issued, industry, and rating; some papers additionally use market-related proxies such as the bid-ask spread, trade volume, number of trades and number of dealers.[4] In the more recent literature, indirect estimators of transaction cost, market impact and turnover have been proposed to analyze liquidity.[5] In the empirical section of this paper, we potentially contribute to this literature by presenting a new measure of liquidity, and by showing its relation to well-established liquidity proxies.

---

[3]See Mahanti et al. (2008) for details of a cross-sectional comparison for the US corporate bond market.

[4]See Elton et al. (2001), Collin-Dufresne et al. (2001), Houweling et al. (2003), Perraudin and Taylor (2003), Eom et al. (2004), Liu et al. (2004), Longstaff et al. (2005), and De Jong and Driessen (2006).

[5]See Amihud (2002), Edwards et al. (2007), Chen et al. (2007), and Mahanti et al. (2008), for example.

Overall, we hope that our proposed market microstructure model based on the determinants of liquidity, together with the empirical results, based on the unique data sets we employ, will improve our understanding of liquidity effects on prices in a relatively illiquid OTC market. Thus, our results are likely to be relevant for many applications in OTC markets, in general, and US fixed income markets, in particular, from the viewpoint of academic researchers, as well as practioners and regulators.

This paper is organized as follows. In Section 2, we develop our market microstructure model for OTC markets and derive our liquidity measure. Section 3 introduces our US corporate bond market data set and presents our results at the level of the market and individual bonds. Section 4 concludes the paper.

2 Model of Price Dispersion

We model a competitive market consisting of I assets and J dealers. The dealers face inventory costs and quote bid and ask prices depending on their desired inventory levels, taking into account the cost of holding inventory, as well as other costs such as those associated with asymmetric information relative to informed investors, and fixed costs of trading. Several investors, who have exogenously given buying or selling needs, trade with the dealers. The market is over-the-counter in nature, implying that an investor has to directly contact dealers to observe their price quotes. In addition, investors face search costs every time they contact a dealer, before they can trade. We model the decision problem for the dealers and for the investors separately.

2.1 Dealers

We assume that there are I assets in the market, each indexed by the identifier $i = 1$ to I, and J dealers, each indexed by the identifier $j = 1$ to J. The inventories of asset i across dealers are randomly distributed. We denote by the quantity $s_{i,j}$ the inventory of asset i with dealer j, which can be positive (long) or negative (short). Each dealer faces an inventory holding cost function H that is convex in the *absolute* quantity of inventory held, and is given by $H = H(s)$, which includes cost of financing the position, as well as implicit costs due to dealers' capital constraints and risk aversion. The marginal holding cost of adding one (infinitesimal) unit of the asset to the inventory is approximated by $h = h(s) = H'(s)$, assuming that the function $H(s)$ is differentiable in s, and has to be considered when trading the asset.

Dealers quote bid and ask prices for a market lot of one (infinitesimal) unit depending on their inventory position, where the ask price of asset i quoted by dealer j is denoted by $p_{i,j}^a$ and the bid price is denoted

by $p_{i,j}^b$.[6] Each dealer takes the marginal holding cost and all other costs arising due to market frictions not related to inventory into account when quoting bid and ask prices. Suppose, that the valuation of dealer j for asset i is given by $m_{i,j}$; then, the outcome of considering these costs is the spread added to the dealer's valuation for computing the ask price and the spread subtracted from this valuation for the bid price. Denoting by f^a and f^b the transaction cost functions which transforms the relevant cost components into these spreads, then the bid and ask prices can be written as

$$p_{i,j}^a = m_{i,j} + f^a(h(s_{i,j})) \tag{1}$$

$$p_{i,j}^b = m_{i,j} - f^b(h(s_{i,j})) \tag{2}$$

Thus, the functions f^a and f^b can be interpreted as the transaction costs faced by an investor when buying or selling from a particular dealer relative to the dealer's valuation. Note that the bid and ask quotes may be asymmetric around the valuation of the asset as the dealer has preferences to buy or sell given his actual inventory. For instance, suppose a dealer has a long position in the asset in question, then this dealer might be more willing to sell bonds than buy bonds due to the inventory costs, i.e., $f^b > f^a$ in this case, since he is more reluctant to increase his inventory. Turning to the market as a whole, we define the market's aggregate valuation of asset i by m_i:

$$m_i = E(m_{i,j}) \quad i = 1\,to\,I, \; j = 1\,to\,J \tag{3}$$

We introduce this notation, as we base our price dispersion measure on the deviations of transacted prices from the market valuation (see Section 2.3).

2.2   Investors

We now consider the problem from the point of view of an investor wishing to execute a trade consisting of a market lot of one (infinitesimal) unit. Since all dealers are identical from the point of view of the investor, let us assume that the dispersion of ask prices faced by an investor wishing to buy one unit of asset i is given by the density function $g_i^a(p^a)$ where p^a is the ask quote when contacting an arbitrary dealer. The investor has already contacted one dealer and is then evaluating the marginal cost and marginal benefit of contacting an additional dealer. Let c indicate the cost of searching for another

---

[6]Note that we assume a market lot of one (infinitesimally small) unit across all dealers for simplicity. All the results presented hold when we allow for differences in the desired lots across dealers, although the results are somewhat more complex.

dealer, and $p^{a,0}$ be the ask price quoted by the dealer with whom the investor is already in contact. Following Garbade and Silber (1976), it can be shown that the investor buys the asset at price $p^{a,0}$ if:

$$p^{a,0} \leq p^{a*} \tag{4}$$

where p^{a*} is a reservation price which solves

$$c = \int_0^{p^{a*}} (p^{a*} - x)g^a(x)dx \tag{5}$$

Similarly, it can be shown that an investor wishing to sell the asset sells it if the bid price is greater than a reservation price that solves:

$$c = \int_{p^{b*}}^0 (x - p^{b*})g^b(x)dx \tag{6}$$

Thus, the investor only contacts an additional dealer, if the expected improvement in the offered price of the bond is higher than the search cost.

2.3 Price Dispersion in Equilibrium

We now proceed to parameterize the problem and draw explicit solutions for the dispersion of transacted prices for specific assumptions for the inventory distribution across dealers, the marginal holding cost function, and the transaction cost function. Let us assume that the inventory of each dealer for asset i is given by the same uniform distribution with support $[\underline{s_i}; \overline{s_i}]$.[7] Furthermore, let us assume that inventory holding costs are independent across assets. This implies that the dealer solves the inventory holding problem for each asset independently ignoring any cross-asset inventory effects. Thus, for any asset, the inventory distribution is uniform for a dealer and is given by $[\underline{s}; \overline{s}]$, suppressing the subscript i.

For simplicity, we assume that the inventory holdings are distributed with a mean of zero, i.e., $-\underline{s} = \overline{s}$. Thus, we assume that the expected net inventory across dealers is zero. For most markets, however, dealers' net inventory is likely be positive if the asset is in positive net supply. In such markets, our assumption can be interpreted in the following way: the dealer's inventory can be separated into two parts, a strategic position and an inventory position attributable to the broker-dealer function. The first part of the dealer's holding can be assumed to be derived from a portfolio optimization decision. The

---

[7]The specific results derived below are dependent on this assumption. However, similar, but perhaps more complex, results can be derived for alternative distribution assumptions.

setup presented here models only the second part, which is assumed to be in zero net supply.

Let us assume that the holding cost for inventory s is given by the convex function

$$H = \frac{\alpha s^2}{4} \tag{7}$$

where α is a positive constant which takes into account all relevant holding costs. This functional form makes inventory holding costs symmetric, whether the inventory is held long or short. It also reflects the increasing reluctance of dealers to hold inventories that deviate substantially from zero, reflecting both capital constraints and risk aversion. This function implies that the marginal holding cost is linear in the inventory holding of the dealer, and is given by:

$$h = \frac{\alpha s}{2} \tag{8}$$

Further, let us assume that the transaction cost functions of each dealer based on all relevant costs are given by

$$f^a = \gamma - h(s) \tag{9}$$

$$f^b = \gamma + h(s) \tag{10}$$

In equation (9) and (10), the transaction costs are modeled as consisting of two parts. The first part, i.e., γ, reflects the non-inventory costs of transacting one unit, including fixed costs of trading and costs due to market frictions. The second part represents the marginal holding costs. Note, that the difference between f^a and f^b is due to the sign of the marginal costs, as trading on the ask side results in an inventory change of minus one unit whereas a bid side trade results in a change of plus one unit from the perspective of the dealer.

For convenience, we scale γ in our model, such that the lower bound of the transaction cost functions is zero. This is achieved by setting $\gamma = \alpha \bar{s}/2$. Given this setup, the transaction cost functions result in identical values for s equal to zero, i.e., the bid and ask quotes are symmetric around the dealer's valuation of the asset for a zero inventory position (see equations 1 and 2). For the largest possible long position, i.e., $s = \bar{s}$, the transaction cost function f^a is zero and $f^b = \alpha \bar{s}$ representing the preference of

the dealer to sell bonds (and vice versa for the largest possible short position). In general,

$$f^a = \frac{\alpha s}{2}(\bar{s} - s) \tag{11}$$

$$f^b = \frac{\alpha s}{2}(\bar{s} + s) \tag{12}$$

In this setup, the transactions costs across dealers are uniformly distributed over $[0; \alpha\bar{s}]$ for ask and bid quotes, given the uniform distribution of the inventory and the transaction cost functions. Hence, there exist dealers who trade at their valuation $m_{i,j}$ of the asset on a set of measure zero. Further, if we assume that all dealers agree on the market's expectation of the price, then by the definition of $m_{i,j} = m_i$, this implies that ask prices are uniformly distributed with support $[m; m + \alpha\bar{s}]$ and that bid prices are uniformly distributed with support $[m - \alpha\bar{s}; m]$ (where we again suppress the asset index i).

Now consider the problem from the point of view of an investor wishing to buy one unit of the security and facing a search cost of c for contacting an additional dealer. Such an investor transacts with dealer j, if the ask quote $p_j^a \leq p^{a*}$ where

$$c = \int_m^{p^{a*}} (p^{a*} - x) g^a(x) dx = \int_m^{p^{a*}} \frac{(p^{a*} - x)}{\alpha\bar{s}} dx \tag{13}$$

Solving the integral gives us the following relationship:

$$p^{a*} = m + \sqrt{2c\alpha\bar{s}} \tag{14}$$

Similarly, solving the equation for an investor wishing to sell one unit of the security gives us the following relationship between the reservation bid price, the mean valuation m, and the search and inventory costs:

$$p^{b*} = m - \sqrt{2c\alpha\bar{s}} \tag{15}$$

Thus, the reservation price for buying (selling) is higher (lower) if the search cost c is high or the cost of inventory holding for an asset is high or inventory is more dispersed across dealers. However, in the absence of market frictions, all trades would take place only at the market's valuation m. With frictions, transactions take place if the offered ask price is less than the reservation ask price, or if the bid price is greater than the reservation bid price. In this setup, two different intervals for the transaction prices are possible depending on the level of the search costs: If the search cost is sufficiently low (i.e., $c \leq \alpha\bar{s}/2$), transactions are distributed over $[p^{b*}; p^{a*}]$, i.e., reservation prices restrict the set of offered

prices. If the search cost is higher than this critical value, investors will accept any prices offer by dealers and transactions will be distributed over the interval $[m - \alpha\overline{s}; m + \alpha\overline{s}]$. In either case, market frictions determine the magnitude of potential deviation of the transaction price from the market's valuation, i.e., they determine the price dispersion of the asset.

Assuming that the appearance of investors wishing to buy and sell the asset is equally likely, transactions are uniformly distributed over this interval. In such a situation, the actual measure of the price dispersion is the variance of transacted prices p_k around the market's expectation of the price m given by:[8]

$$E(p_k - m)^2 = \begin{cases} \frac{2}{3}c\alpha\overline{s} & \text{if } c \leq \alpha\overline{s}/2 \\ \frac{1}{3}\alpha^2\overline{s}^2 & \text{otherwise} \end{cases} \tag{16}$$

Again, this variance is a function of the market frictions, i.e., increasing in the search cost of the investor c (if the reservation prices are binding, i.e., $c \leq \alpha\overline{s}/2$), the inventory cost and the distribution of inventories across dealers for that asset.[9] Note that the variance (or volatility) of the price dispersion can be estimated, if the transaction prices and the respective valuation of the market are available. Thus, this derivation of the price dispersion variance is the main result of our model, which we use to define a new liquidity measure in Section 2.4.

In some markets, the average or median bid-ask quotes across all dealers are made public, e.g. at the end of the trading day. Often it is assumed that these quotes represent bounds, within which most of the trades take place. Much discussion is ongoing whether this is a realistic assumption. Our model can provide an analytic solution to explore this question. We define the hit-rate along the lines of Bliss (1997), who calculates the percentage of cases where a certain price lies within the range spanned by bid and ask quotations. We define the hit-rate HR as the percentage of trades that fall within the median bid-ask quote; then, this percentage in our model is represented by the probability that a traded price in the possible range lies within the mean or median bid-ask quote represented by the range $[m - \alpha\overline{s}/2; m + \alpha\overline{s}/2]$. Three different ranges for the hit-rat are possible depending on the level of the search costs: If the search cost are lower than $\alpha\overline{s}/8$, then no prices above the median bid-ask quote are accepted by the investors and the hit-rate is 100%. If the search costs are greater than $\alpha\overline{s}/2$, then all quotes are accepted by investors and the hit-rate is 50% by construction. For intermediate search costs, the hit-rate depends on

---

[8]This results follows directly from the functional form of the variance for an uniformly distributed random variable.

[9]Although the explicit functional form of the price dispersion measure derived above depends on the uniform distribution for dealers' inventory, the general result that it depends on search costs and the marginal holding costs will always hold in this framework.

the market friction parameters:

$$HR = \begin{cases} 50\% & \text{if } c > \alpha\bar{s}/2 \\ \frac{\alpha\bar{s}}{2\sqrt{2c\alpha\bar{s}}} = \frac{\sqrt{\alpha\bar{s}}}{2\sqrt{2c}} & \text{if } \alpha\bar{s}/8 \leq c \leq \alpha\bar{s}/2 \\ 100\% & \text{if } c < \alpha\bar{s}/8 \end{cases} \qquad (17)$$

This implies that the hit rate increases when the cost of searching c is lower, or when the inventory cost and the dispersion of quoted spreads given by $\alpha\bar{s}$ is higher. It is also clear that there is no reason for the hit rate to be close to 100%, as this is determined by the market frictions. In fact, in general, the hit rate depends on both how dispersed quotes are, and how costly it is to search for a new dealer. When quotes are dispersed, and it is costly to search for new dealers, transacted prices may be regularly outside the mean or median bid-ask spread observed in the market.

2.4 Liquidity Measure

Based on the framework presented in the previous sections, we propose the following new liquidity measure to quantify the price dispersion per bond on a daily basis. The measure is based on the transaction prices and volumes, and on the respective market's expectation of the price: On each day t, for bond i, we observe $K_{i,t}$ traded prices $p_{i,k,t}$ (for $k = 1$ to $K_{i,t}$) and one market-wide valuation $m_{i,t}$. Each traded price has a trade volume of $v_{i,k,t}$. Based on this information, we define the new liquidity measure $d_{i,t}$ as

$$d_{i,t} = \sqrt{\frac{1}{\sum_{k=1}^{K_{i,t}} v_{i,k,t}} \cdot \sum_{k=1}^{K_{i,t}} (p_{i,k,t} - m_{i,t})^2 \cdot v_{i,k,t}} \qquad (18)$$

This measure represents the root mean squared difference between the traded prices and the respective market-wide valuation based on a volume-weighted calculation of the difference. Thus, this measure is an estimate of the absolute deviation, and, more importantly, has the interpretation as the volatility of the price dispersion, as derived in equation (16). Therefore, this measure, which can be thought of as the reduced form of the model presented earlier, is of empirical interest, when analyzing price dispersion effects.[10]

We use a volume-weighted difference measure since we assume that price dispersions in larger trades

---

[10]Using the model presented, this measure could be calculated based on the investors search costs, the distribution of inventory across dealers, and the marginal holding cost, see equation (16). However, this information is in general not accessible.

reveal more reliable deviations from the market's average valuation. Furthermore, this weighting can be seen as a device for the elimination of outliers of potentially erratic prices for particulary small trades. Alternatively, for such trades, we could have excluded trades below a certain trade size. However, this may not be appropriate as the average trade size can vary significantly across bonds.

Note that, in this calculation, we implicitly assume that the difference between the traded prices and the market-wide (end-of-day) valuations is not influenced by the trading time during a particular day. In other words, we assume that all transactions occurring in a given day arise at the same time. Given the infrequency of trades in the corporate bond market, this is not an unreasonable approximation. However, we analyze the robustness of our measure with respect to changes in this assumption in Section 3.2.3.

3   Empirical Analysis

According to the model presented in Section 2, inventory risk and search costs determine the probability distribution of prices for buy and sell transactions of investors in OTC markets. Hence, in our framework, illiquidity is interpreted as the potential cost of trading in the presence of these frictions. The less liquid an asset, the more likely is a significant deviation of the actual observed transaction price from the market's expectation of the price. In our empirical example, we analyze the liquidity of the US corporate bond market. This market is an important and well-known financial market, and especially interesting for our purposes, since liquidity differences between individual bonds appear to be rather pronounced.

We quantify our liquidity measure and analyze the price dispersion effects at the market and individual bond level. In particular, we relate our measure to conventional liquidity proxies to confirm that it indeed represents liquidity. Furthermore, we compare our liquidity measure with bid-ask spreads quoted on Bloomberg to allow for an economic interpretation of the results. This is especially of interest, as bid-ask spreads themselves are generally regarded as proxies for price dispersion effects as well, and are often used to measure liquidity effects, since transaction data are rarely available in OTC markets. By exploring the actual hit-rate for the data set, our analysis may point to the validity of using the bid-ask spread as a liquidity metric.

3.1   Data Description

The US corporate bond market offers a unique opportunity to calibrate and test market microstructure models. Unlike other OTC or dealer markets, a central data source exists for all transactions in this

market. The National Association of Securities Dealers (NASD), now known as the Financial Industry Regulatory Authority (FINRA), established the Trade Reporting and Compliance Engine (TRACE) in October 2004, making the reporting of all transactions in US corporate bonds obligatory for all brokers/dealers under a set of rules approved by the Securities and Exchange Commission (SEC). TRACE reporting by broker-dealers was introduced in three consecutive phases.[11] Phase I started in July 2002 and covered only reporting for the larger and generally higher-credit quality issues. Phase II expanded the dissemination to smaller investment grade issues. Phase III started on October 1, 2004 and reporting then covered all secondary market transactions for corporate bonds. As a result of the TRACE initiative, we can obtain all transaction prices and volumes for this market, whereas in other OTC markets, this information has either to be approximated by using transaction data from only one or a small set of dealers, or by using bid-ask quotations instead.

Besides the transaction data, a second important source of valuation/mark-to-market information exists. This data set is provided by Markit Group Limited, which was founded in 2001 as a private company. One of its services is to collect, validate, aggregate, and distribute end-of-day composite bond prices, where the input information is collected from more than thirty major dealers in the market, who provide price information from their books and from automated trading systems.[12] Various data cleaning and aggregation procedures are applied, and thus, the resulting Markit quotations can be interpreted as a market-wide average or expectation of the price of a particular bond.[13] Markit quotations are publicly available for a fee and are used by many financial institutions as the main price information source to mark their portfolios to market, since they are seen as more reliable than end-of-day bid-ask quotations. Combining TRACE prices and Markit quotations allows us to calculate our liquidity measure for the US corporate bond market.

Our data set consists of TRACE prices, Markit quotations, and Bloomberg bid-ask quotations (close ask / close bid) available for the time period October 1, 2004 to October 31, 2006. This period starts just after the implementation of Phase III of the TRACE project, when all secondary market transactions were reported to the database. For our analysis, we include only coupon- and floating-rate dollar denominated bonds with a bullet or callable repayment structure, without any other option features, which were traded on at least 20 days in the two year period. Furthermore, we restrict our sample to bonds for which issue ratings from Standard & Poor's, Moody's, or Fitch are available, to include credit information, facilitating

---

[11]See National Association of Securities Dealers (2006).
[12]See Markit (2006).
[13]Ibid.

the study of the correlation between credit risk and liquidity in our analysis. Even with these restrictions, the data sets results in 1,800 bonds with 440,076 Markit/Bloomberg quotations and 3,889,017 TRACE prices. Note that there is only one data point per bond and day available from Markit and Bloomberg, whereas typically several transactions are reported in TRACE for the more liquid bonds.

As a result of the screening criteria we use, the selected bonds represent 7.98% of all corporate bonds available in TRACE, i.e., of all bonds that had at least one trade in TRACE in the observed time period. However, our data set accounts for an amount outstanding of \$1.308 trillion, which represents 25.31% of the total amount outstanding of all bonds as on June 30, 2006.[14] Based on the share of trading activity, the selected sample represents even a higher proportion, accounting for 37.12% of the total trading volume. Thus, our data set is representative of the US corporate bond market, with the advantage that each bond in the sample has sufficient observations along with important additional variables for empirical analysis.

Overall, the selected bonds represent an important segment of the corporate bond market with a slight bias toward those with high liquidity compared to the rest of the market. As Edwards et al. (2007) report for their TRACE sample from 2003, only 16,746 bonds out of almost 70,000 have more than 9 trades per year and Mahanti et al. (2008) report that over 40% of the bonds in their sample, do not even trade once a year. However, even this selected segment has rather low overall liquidity. To emphasis this point, Table 1 shows the trading frequency of our bond sample measured by the number of trading days for the two available one year periods, i.e., 10/2004 to 10/2005 and 10/2005 to 10/2006, respectively.

[Table 1 around here.]

The bonds are divided into five equidistant categories, i.e., traded on up to 50, 51 to 100, 101 to 150, 151 to 200, and on more than 200 days per year. Table 1 shows that the bonds are nearly evenly distributed over the categories with a slightly higher concentration in the lowest category, i.e., 27.39% of all bonds in the first one-year period and 26.94% in the second period show very low trading activity with less than 50 trades per year indicating low overall liquidity.

For a bond-level analysis of these liquidity effects, we add bond characteristics and trading activity variables to our data set. The bond characteristics contained in our data set include coupon, maturity, age, amount issued, issue rating, and industry, and the trading activity variables are trade volume, number of trades, bid-ask spread, and depth (i.e., number of major dealers providing a quote to Markit) per bond. The issue rating represents the actual rating of a bond as on October 1, 2007 obtained from

---

[14]The amount of bonds outstanding on June 30, 2006 was \$5.167 trillion, based on data from the Bond Market Association, supplied by Reuters.

Bloomberg (or, for matured bonds, the last valid rating), and therefore is a rough proxy for the rating at the end of the selected time period. For all other variables we have the complete time-series available.[15]

[Figure 1 and Figure 2 around here.]

Figure 1 and 2 show the distribution of bonds across industries and credit rating grades, respectively. The industry categories in Figure 1 are obtained from Bloomberg. The ratings in Figure 2 represent the "average" issue rating per bond from Standard & Poor's, Moody's, and Fitch, generated by first transforming the individual ratings to numerical values (AAA=1 to CCC=7) and then using the rounded mean. The industry distribution shows the expected result that the sample has higher concentrations in the banking/financial and the industrial sectors. The rating distribution is also skewed, with a higher concentration in investment grades bonds, especially in the A and BBB ratings. These numeric values are directly used in the regression analysis, presented in the next section.[16]

[Table 2 around here.]

Table 2 shows the distribution of all the other variables at the bond level, where time-variant variables are represented by their time-series averages. For the trading activity variables, the average is computed by averaging the values for days on which the bond was actually traded, i.e., the average is conditional on trading. The coupons and bid-ask spreads of the bonds are in the ranges of 1.95% to 11.25% with a median of 6.125%, and 1.55 bp to 90 bp with a median of 32.17 bp, respectively. The maturity spectrum shows a concentration of bonds with a maturity of up to ten years. The amount issued is quite dispersed going from $100 million up to $6.5 billion with a median of $500 million. The average trade volume per day goes from just over $100,000 up to $61 million with an median of $2.8 million. The average number of trades per day goes from one trade to 121 trades with a median of 4 trades. The depth of the Markit quotations shows that, on average, quotes come from five dealers with a minimum of three and a maximum of thirteen.[17] Overall, these summary statistics indicate that the chosen portfolio is representative and comparable to data sets used by prior studies on this market.[18]

---

[15]For floating-rate bonds, we have the time-series of actual coupon rates available.

[16]Alternatively, the ratings could have been represented by their probability of default (PD) equivalents. However, since each rating agency provides different PD estimates and does not clearly state the specific time horizon over which the ratings are calibrated (and whether ratings are at all calibrated on PDs), we do not apply such procedures.

[17]Note that Markit collects data from more than thirty dealers, but not all dealers provide information for each bond on every day.

[18]See e.g. Edwards et al. (2007) and Mahanti et al. (2008).

3.2 Results

3.2.1 Market Level Analysis

We are mainly interested in the deviations between TRACE prices and Markit quotations when analyzing liquidity based on price dispersion effects in the context of the US corporate bond market. Assuming that the Markit quotation represents the mean of the price dispersion distribution, we would not expect that the average deviation (where positive and negative differences cancel out), in the market as a whole would be biased in either a positive or negative direction. In our data set, this volume-weighted average difference between TRACE prices and respective Markit quotations is 4.88 bp with a standard deviation of 71.85 bp, i.e., on average, we find that the Markit composite is slightly lower than the TRACE price: However, the difference does not indicate an economically significant bias, given that the average quoted bid-ask spread on Bloomberg is 35.90 bp in this market.

Turning to the analysis of price dispersion effects, we calculate our proposed liquidity measure for each bond on a daily basis. The measure represents the root mean squared difference between the TRACE prices and the respective Markit quotation. To analyze the liquidity of the corporate bond market as a whole, we calculate the average of this liquidity measure over all bonds and all days. The resulting average root mean squared difference is 49.94 bp and the standard deviation of this liquidity measure is 63.36 bp. This large variation is mainly driven by the cross-sectional differences across bonds: taking the time-series average of the measure for each bond, we find a minimum of 1.99 bp and a maximum as high as 351.90 bp (with 10% quantile of 16.42 bp and 90% quantile of 97.99 bp). Overall, we find large differences between traded and expected prices, indicating a rather low overall liquidity.

As expected, the liquidity measure for the whole corporate bond market is larger than the average quoted bid-ask spread of 35.90 bp. A t-test for equal means reveals that the mean of our measure is significantly higher than that observed for the bid-ask spread. Even more pronounced than the difference in the means is the difference in the standard deviation which is nearly three times larger, as the standard deviation of the bid-ask spreads is 23.73 bp. This difference can also be illustrated by the following statistic: our liquidity measure is larger than the quoted bid-ask spread in 50.59% of all cases, i.e., we find that bid-ask quotations can only roughly approximate liquidity costs.

As mentioned in Section 2.4, in the calculation of our liquidity measure, we implicitly assume that the difference between the traded prices and the market-wide end-of-day valuation is not influenced by the

trading time.[19] This is a strong assumption. However, in our data set we cannot find any time dependence during the trading day. As a straight-forward test at the level of the market, we calculated our measure using prices before 1 pm and after 1 pm, respectively, as 1 pm is the median trade time. We find, that the difference is statistically insignificant and even has the wrong sign, i.e., the price dispersion measure is higher for prices closer to the end of the trading day. The result stays the same when we compare the measures calculated with prices from two separated time spans, i.e., before 12 pm and after 2 pm, as a stress test.

Summing up the market level analysis, we find significant differences between TRACE prices and Markit composite quotations that cannot be simply explained by bid-ask spreads or trade time effects. The model presented earlier relates these differences to liquidity. The large variation of our liquidity measure across bonds strengthens the liquidity argument as a quick analysis also shows significant differences in the trading frequency, as documented in Section 3.1. The question arises whether the hypothesis that our measure represents liquidity can be further supported by bond level analysis.

3.2.2   Bond Level Analysis

In the analysis at the individual bond level, we relate our liquidity measure to bond characteristics and trading activity variables, which have been used or proposed as liquidity proxies by other researchers. A strong relation between these variables and the new measure would support our liquidity argument. As bond characteristics, we use the coupon, maturity, log(amount issued), age, rating and industry. As proxies for trading activity we employ log(trade volume) in TRACE, number of trades in TRACE, number of reporting dealers in Markit, and the bid-ask spread on Bloomberg. Table 3 presents the correlations between the time-weighted averages of these variables and the price dispersion measure across bonds.

[Table 3 around here.]

The table shows high correlations between almost all variables and our liquidity measure. Figure 3 shows this relation graphically for six variables.

[Figure 3 around here.]

To formally test for a significant relationship between our measure and the other variables, we employ cross-sectional linear regressions, again using time-weighted averages of the variables. We present results using time-averages based on the whole time period, as well as based on each available quarter (2004 Q4

---

[19]This assumption is necessary as we have only one bond-specific market-wide valuation per day available.

to 2006 Q3), to check whether time-series effects drive the results and have to be considered explicitly. Univariate regressions (not presented in detail) reveal that basically every variable has a statistically significant relation with our measure, whereas the R^2s vary considerably.[20] We use the univariate results to build a multivariate regression model explaining the observed bond-specific liquidity measure. Table 4 presents the resulting coefficient estimates.

[Table 4 around here.]

The results for the multivariate regression for the whole time period shows the high explanatory power of the model represented by an R^2 of 61.5%. Considering the regressions for the individual quarters, we again find high values for the R^2 varying between 44.9% and 51.3%. These numbers are impressively high and extremely stable over time showing a strong relation between our measure and liquidity-related covariates. The variables turning out to be significant in the multivariate model are maturity, amount issued, age, rating, bid-ask spread, and trade volume, i.e., the price dispersion measure shows a strong relation to bond characteristics, as well as to trading activity. Their influence and statistical significance stays surprisingly stable over time. All the variables have the expected sign, as explained in the next five paragraphs.

To illustrate the influence of the different variables, we show the effect on our liquidity measure caused by a one standard deviation change in a particular explanatory variable, as calculated from the available set of bonds. The variable with the strongest effect is the actual maturity of the bond. The liquidity measure of bonds with a one standard deviation greater maturity than other bonds (i.e., longer maturity of 7.11 years) is expected to be 13.09 bp higher, i.e., the higher the maturity the lower the underlying liquidity of the bond. Analyzing our data set in more detail, we find that bandwidth maturities between zero and ten years show minor effects on the liquidity measure. The regression parameter seems to be driven by the lower liquidity of bonds with maturities of more than ten years (see Figure 3). This is an expected result as these long-term bonds are mainly held by "buy-and-hold" investors such as pension-funds and insurance companies, which match the duration of the bonds they hold with that of their long term liabilities. Consequentially, the liquidity of these bonds is expected to be low.

The bid-ask spread has the second strongest effect. As expected, we find a positive correlation with our price dispersion measure (see Figure 3). A one standard deviation increase in the bid-ask spread (i.e. 20.23 bp higher bid-ask spreads) results in a 11.49 bp increase in the price dispersion measure. This

---

[20]The number of reporting dealers in Markit has the lowest univariate explanatory power represented by an R^2 of 3.57% and the bid-ask spread has the highest explanatory power represented by an R^2 of 39.96%.

indicates that the bid-ask spread can roughly proxy for the effective costs of inventory risk and search costs, and thus, has a relation to liquidity as argued in many papers. However, the magnitude may not exactly be in line with the liquidity costs, as we would expect a one-to-one relation in this case.

We find the expected negative correlation between trade volume and the liquidity measure. The effect is -6.88 bp for a one standard deviation increase in log(trade volume). This parameter shows that variables that are directly related to trading activity have a clear influence on the liquidity measure. However, the influence is lower than expected, given that other characteristics have a stronger influence. This result shows that trading activity itself is an important aspect of liquidity, but it is not the only variable determining the price impact for instantaneously trading large volumes.

A positive correlation can be found for the age of a bond. A bond with a greater age has a lower liquidity (one standard deviation, i.e., 2.87 years, results in an increase in the price dispersion of 5.93 bp). This can be attributed to the well-known effect of "on-the-run" and "off-the-run" bonds, where much higher liquidity is available for newly issued bonds (i.e., "on-the-run" bonds) compared to the liquidity after the bonds have been outstanding for some time (i.e., "off-the-run" bonds).

Our results indicate lower liquidity for lower credit quality bonds as is assumed by most market participants (one standard deviation, i.e., approximately one rating grade, results in an increase in the price dispersion of 4.67 bp). Analyzing our data set, we further find that the results of the credit quality variable are mainly driven by the difference between investment grade and speculative grade bonds (see Figure 3). As expected, liquidity is higher for the investment grade segment.[21] Based on the resulting regression parameter, the difference in the liquidity measure between an average investment grade and speculative grade bond (represented by a difference of four ratings grades) is 18.68 bp.

The variable in the multivariate model with the least influence is the amount issued for a bond. The higher the log(amount issued), the higher the liquidity. A one standard deviation increase has an effect of -2.15 bp on the price dispersion. This is the expected result as many papers claim that a certain volume can be traded more easily within a large issue than within a small issue, although the magnitude of the effect is quite small.

Overall, we find a strong relationship between our liquidity measure, and bond characteristics and trading activity. Therefore, we see the liquidity interpretation of our proposed measure confirmed, as it is related to many variables cited as liquidity proxies by other researchers.

---

[21]Note that we only have the rating information at one point in time available, i.e., we cannot detect the time-series effects of potentially higher liquidity for bonds that are getting close to a rating change, especially on the cusp between investment and speculative grade. See Mahanti et al. (2008) for a discussion.

To further validate this result, we analyze the explanatory power of our liquidity measure in predicting established estimators of liquidity in the literature. One important approach to define liquidity is through the price impact of trading a particular financial instrument. Several measures based on the price impact have been proposed in the microstructure literature. The measure introduced by Amihud (2002), indirectly based on the concept introduced by Kyle (1985), is most appropriate for this purpose and widely accepted in the literature. The Amihud measure is the average ratio of the absolute return of a particular bond to its trading volume for a given period. Defining the absolute return of a bond on day t by $|r_t|$, and the trading volume by v_t, the Amihud measure for a time period of T days is given as

$$\text{Amihud measure} = \frac{1}{T} \sum_{t=1}^{T} \frac{|r_t|}{v_t} \tag{19}$$

We use the volume-weighted average TRACE price each day to generate the returns. Days on which there are no trades represent zero return and zero trading volume and are thus not included in the average. For the calculation of the measure for quarterly intervals, we use the restriction that bonds have to be traded at least on five days to avoid erratic values for the measure.

A high Amihud measure implies a larger price impact of trading, and therefore, low liquidity. To analyze the relation between the Amihud measure and our liquidity measure, we first use a univariate regression where the log(Amihud measure) is the dependent variable. Again, we apply a cross-sectional regression by calculating time-series averages for the whole time period and the individual quarters. Table 5 shows the results of these regressions.

[Table 5 around here.]

The results show the expected positive correlation of the Amihud measure with our liquidity measure. The parameter is statistically significant for all time periods with the expected positive sign. The R^2 is high for all observed time periods (between 19.8% and 31.3%). Again, the regression parameter representing our liquidity measure stays surprisingly stable over the quarters. The parameter also represents an economically significant effect. In an overall sense, a one standard deviation increase in our measure leads to a 300% increase in the Amihud measure.

To analyze whether our proposed measure has explanatory power, over and above other liquidity proxies, in explaining the Amihud measure, we apply a multivariate regression approach additionally using all available bond characteristics and trading activity variables. Results of the multivariate regression model are presented in Table 6.

21

[Table 6 around here.]

The resulting model has high values for R^2 (42.5% to 63.7%), with our liquidity measure exhibiting an economically and statistically significant parameter for all time periods. The size of the parameter is comparable to that of the univariate model. Most of the other significant parameters are related directly or indirectly to trading activity, i.e., number of trades, log(trade volume), and log(amount issued) with the expected negative sign. When we drop our liquidity measure from the regression as explanatory variable, we find a significant decrease in R^2 in the range of eight to ten percent points indicating the high explanatory power over and above the other covariates. This validation strengthens the interpretation of our liquidity argument.

3.2.3   Hit-Rate Analysis of the Bid-Ask Quotations

Although the price dispersion and its analysis, at the market-wide and bond-specific level, is of main interest for our research question, it is worthwhile to compare TRACE prices and Markit composites directly with quoted bid and ask quotations. In many studies, bid-ask quotations (or mid quotations) have been considered as proxies for traded prices since such traded prices were not available for the corporate bond market in prior years (and are still not available for most other OTC markets). These studies usually assume that most trades take place within the bounds of the bid-ask quotes and that these quotations are, therefore, representative. Given our data set for the corporate bond market, we have the opportunity to evaluate this assumption by quantifying the hit-rate. Our model presented in Section 2 derives an analytic solution for the hit-rate and shows that the actual ratio depends on the search costs and the inventory costs, and thus, that the actual hit-rate will not be necessarily close to 100%, as shown in equation (17).

Turning to the data, the volume-weighted hit-rate for the TRACE prices is 51.12% (i.e., in these cases, the TRACE price lies within the bid and ask range quoted on Bloomberg). The deviations are roughly symmetric, as out of the misses, 50.12% are lower than the bid prices, and 49.88% are higher then the ask prices, respectively. These results indicate that deviations of traded prices from bid-ask quotations are far more frequent than assumed by most studies. Thus, caution has to be exercised when drawing conclusions in OTC markets, based on bid-ask quotations. The hit-rate for the Markit composite with a value of 58.59% is somewhat higher, but surprisingly low as our model assumes a hit-rate of 100% for this comparison. As there is no theoretical reason for a deviation, we expect that this result indicates

22

a low data quality of the bid-ask quotations in general, e.g. caused by stale quotes or quotes that hold only for small quantities.

4   Conclusion

In this paper, we present a market microstructure framework, modeling price dispersion effects in over-the-counter (OTC) markets with market frictions. We show that, in the presence of stochastic inventories with costs/restrictions for dealers and search costs for investors, traded prices may differ from average market valuations. We interpret these price dispersions as a liquidity effect and derive a new liquidity measure, based on a metric of the deviations between the traded prices and their respective market valuations.

In our empirical implementation, we show results for the US corporate bond market, which presents a unique opportunity to study this liquidity effect. From October 2004 onwards, all OTC transactions in this market have to be reported to the centralized TRACE database. Additionally, market-wide aggregated valuations are available from Markit Group Limited. The combination of these two sources offers us the opportunity to directly observe price dispersion effects in an OTC market, for the first time.

An analysis at the level of the whole market, as well as at the bond level, shows that according to our new liquidity measure, price dispersion effects are significantly larger than quoted bid-ask spreads. These results indicate that the overall liquidity of the corporate bond market is rather low, and that liquidity, so far, could only be roughly approximated by quotations. Furthermore, we show that our new measure is indeed related to liquidity. For this purpose, the new measure is regressed on commonly-used liquidity proxies, i.e., bond characteristics and trading activity variables, as well as established liquidity measures used in the academic literature. The results show a strong relationship with these proxies. According to our results, the most important liquidity proxies for corporate bonds are the amount issued, maturity, age, rating, bid-ask spread, and trading volume, as well as the price impact measure introduced by Amihud (2002).

In addition, our results serve as an indirect test of the reliability of end-of-day marks provided by using the average prices or bid-ask quotations. Calculating a volume-weighted hit-rate, we find that only 51.12% of the TRACE prices and 58.59% of the Markit quotations lie within the bid and ask range quoted on Bloomberg. These numbers are far smaller than previously assumed.

Thus, our results are relevant for many applications in OTC markets, such as bond pricing, risk man-

agement, bond issuance decisions, investment decisions, portfolio optimization, and financial market regulation. Overall, these findings foster a better understanding of OTC markets by analyzing the theoretical and empirical drivers of liquidity effects and provide a clear interpretation of different price data available in these markets to investors.

References

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time series effects. *Journal of Financial Markets*, 5:31–56.

Amihud, Y. and Mendelson, H. (1980). Dealership market: Market making with inventory. *Journal of Financial Economics*, 8:31–53.

Amihud, Y., Mendelson, H., and Pedersen, L. H. (2006). Liquidity and asset prices. *Foundations and Trends in Finance*, 1:269–364.

Bagehot, W. (1971). The only game in town. *Financial Analysts Journal*, 27:12–22.

Bliss, R. R. (1997). Testing term structure estimation methods. *Advances in Futures and Options Research*, 9:197–232.

Bollen, N. P. B., Smith, T., and Whaley, R. E. (2004). Modeling the bid/ask spread: Measuring the inventory-holding premium. *Journal of Financial Economics*, 72:97–141.

Chen, L., Lesmond, D. A., and Wei, J. (2007). Corporate yield spreads and bond liquidity. *The Journal of Finance*, 62:119–149.

Collin-Dufresne, P., Goldstein, R. S., and Martin, J. S. (2001). The determinants of credit spread changes. *The Journal of Finance*, 56:2177–2207.

De Jong, F. and Driessen, J. (2006). Liquidity risk premia in corporate bond markets. *Working Paper, University of Amsterdam.*

Duffie, D., Garleanu, N., and Pedersen, L. H. (2005). Over the counter markets. *Econometrica*, 73:1815–1847.

Duffie, D., Garleanu, N., and Pedersen, L. H. (2007). Valuation in over the counter markets. *The Review of Financial Studies*, 20:1865–1900.

Easley, D. and O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19:69–90.

Edwards, A., Harris, L., and Piwowar, M. (2007). Corporate bond market transaction costs and transparency. *Journal of Finance*, 62:1421–1451.

Elton, E. J., Gruber, M., Agrawal, D., and Mann, C. (2001). Explaining the rate spread on corporate bonds. *The Journal of Finance*, 56:247–277.

Eom, Y., Helwege, H. J., and Huang, J. Z. (2004). Structural models of corporate bond pricing: An empirical investigation. *The Review of Financial Studies*, 17:499–544.

Garbade, K. D. and Silber, W. L. (1976). Price dispersion in the government securities market. *The Journal of Political Economy*, 84:721–740.

Garman, M. (1976). Market microstructure. *Journal of Financial Economics*, 3:257–275.

Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14:71–100.

Grossman, S. J. and Miller, M. H. (1988). Liquidity and market structure. *The Journal of Finance*, 42:617–633.

Hansch, O., Naik, N. Y., and Viswanathan, S. (1998). Do inventories matter in dealership markets? evidence from the london stock exchange. *The Journal of Finance*, 53:1623–1656.

Ho, T. and Stoll, H. R. (1980). On dealer markets under competition. *The Journal of Finance*, 35:259–267.

Ho, T. and Stoll, H. R. (1983). The dynamics of dealer markets under competition. *The Journal of Finance*, 38:1053–1074.

Houweling, P., Mentink, A., and Vorst, T. (2003). How to measure corporate bond liquidity. *Working Paper, Erasmus University Rotterdam*.

Huang, R. D. and Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *The Review of Financial Studies*, 10:995–1034.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53:1315–1335.

Liu, J., Longstaff, F. A., and Mandell, R. E. (2004). The market price of credit risk: An empirical analysis of interest rate swap spreads. *Journal of Business*, 79:2337–2359.

Longstaff, F., Mithal, S., and Neis, E. (2005). Corporate yield spreads: Default risk or liquidity? new evidence from the credit-default swap market. *The Journal of Finance*, 60:2213–2253.

Mahanti, S., Nashikkar, A., Subrahmanyam, M., Chacko, G., and Mallik, G. (2008). Latent liquidity: a new measure of liquidity, with an application to corporate bonds. *Journal of Financial Economics forthcoming.*

Markit (2006). Markit.com user guide - version 8.0 march 2006. User guide, Markit Group Limited.

National Association of Securities Dealers (2006). Trace fact book - 2006. Fact book, NASD.

Perraudin, W. and Taylor, A. (2003). Liquidity and bond market spreads. *Working Paper, Bank of England.*

Trading Frequency	10/2004 to 10/2005	10/2005 to 10/2006
> 200 days in year	411	392
151-200 days in year	309	369
101-150 days in year	236	322
51-100 days in year	221	222
≤ 50 days in year	444	459
total number of bonds	1621	1704

Table 1: This table shows the trading frequencies of US corporate bonds in our data set, measured by the number of days per year on which a particular bond is traded. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit.

	Coupon (percent)	Maturity (years)	Age (years)	Amount Issued (dollars)	Bid-Ask (basis points)	Trade Volume (dollars)	Trades (number)	Depth (number)
Minimum	1.95	0.07	0.04	100,000,000	1.55	127,925	1.21	3.00
10% Percentile	3.8	1.01	0.70	250,000,000	12.20	1,040,203	2.23	3.29
20% Percentile	4.625	2.22	1.12	300,000,000	18.41	1,510,320	2.62	3.57
30% Percentile	5.2	3.21	1.94	350,000,000	24.04	1,858,653	3.00	3.87
40% Percentile	5.625	4.05	2.53	480,000,000	28.12	2,289,310	3.50	4.20
50% Percentile	6.125	4.80	3.34	500,000,000	32.17	2,797,218	4.14	4.62
60% Percentile	6.55	5.80	4.95	600,000,000	36.48	3,418,796	5.10	5.17
70% Percentile	6.965	6.80	5.59	750,000,000	42.30	4,308,220	6.40	5.81
80% Percentile	7.45	8.89	7.370	1,000,000,000	51.40	5,581,289	8.91	6.87
90% Percentile	7.875	20.79	12.72	1,500,000,000	66.45	7,941,768	14.73	8.26
Maximum	11.25	31.61	16.23	6,500,000,000	90.00	61,274,000	121.60	12.86

Table 2: This table shows the distribution of bond characteristics and trading activity variables within our corporate bond data set. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit. The numbers are presented at the bond-level, i.e., the variables stay either constant over time, e.g. amount issued, or represent times-series averages of the daily data, e.g. average daily trade volume per bond. The table shows the minimum and maximum value for each variable across bonds, as well as the percentiles.

	Price Dispersion	Coupon	Maturity	Age	Amount Issued	Bid-Ask	Trade Volume	Trades	Depth	Rating
Price Dispersion	1.000									
Coupon	0.414	1.000								
Maturity	0.566	0.247	1.000							
Age	0.380	0.527	0.099	1.000						
Amount Issued	-0.265	-0.152	-0.055	-0.166	1.000					
Bid-Ask	0.632	0.495	0.465	0.160	-0.196	1.000				
Trade Volume	-0.186	-0.144	0.098	-0.322	0.566	-0.076	1.000			
Trades	0.028	0.018	-0.020	-0.038	0.590	0.001	0.545	1.000		
Depth	-0.190	-0.061	0.069	-0.321	0.646	-0.088	0.536	0.536	1.000	
Rating	0.429	0.465	0.140	0.211	-0.235	0.584	-0.032	0.029	-0.232	1.000

Table 3: This table shows the correlations of the proposed liquidity measure, bond characteristics, and trading activity covariates across bonds. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit. The liquidity measure is based on the price dispersion and represents the root mean squared difference between TRACE prices and the respective Markit quotation calculated on a daily basis for each bond. TRACE prices are the actual traded prices and the Markit quotation represents the average market valuation, which we interpret as the market's expectation of the price. Thus, the proposed measure represents the volatility of the observed price deviations. The cross-sectional correlations presented are calculated based time-weighted averages of all variables.

	Constant	Maturity	Amount Issued	Age	Rating	Bid-Ask	Trade Volume	R^2	Observations
2004 Q4	231.732	2.576	-5.597	3.849	2.090	0.237	-7.963	0.449	1270
	(7.010***)	(19.781***)	(-2.959***)	(8.795***)	(6.302***)	(3.778***)	(-6.941***)		
2005 Q1	202.072	2.054	-2.537	2.486	1.831	0.490	-10.245	0.469	1417
	(7.159***)	(17.825***)	(-1.575)	(6.688***)		(9.520***)	(-9.911***)		
2005 Q2	221.567	2.692	-5.208	3.560	1.992	0.387	-8.032	0.513	1470
	(7.707***)	(22.764***)	(-3.136***)	(9.446***)	(6.587***)	(7.512***)	(-7.626***)		
2005 Q3	211.069	2.053	-4.268	2.059	1.427	0.360	-8.197	0.461	1517
	(8.263***)	(18.485***)	(-3.048***)	(6.342***)	(5.275***)	(7.361***)	(-8.515***)		
2005 Q4	125.884	1.907	-0.995	2.901	2.214	0.505	-7.794	0.480	1529
	(4.543***)	(16.676***)	(-0.634)	(8.355***)	(7.517***)	(10.182***)	(-7.291***)		
2006 Q1	161.144	1.450	-1.496	1.455	1.464	0.471	-8.799	0.498	1536
	(7.266***)	(15.259***)	(-1.253)	(5.183***)	(6.318***)	(11.877***)	(-10.320***)		
2006 Q2	204.960	1.618	-3.734	1.282	0.983	0.445	-8.442	0.496	1548
	(9.361***)	(16.313***)	(-3.145***)	(4.813***)	(4.245***)	(10.764***)	(-10.026***)		
2006 Q3	167.760	1.453	-3.710	1.242	1.096	0.544	-6.023	0.493	1513
	(7.353***)	(14.254***)	(-3.031***)	(4.509***)	(4.415***)	(12.550***)	(-7.643***)		
Overall	187.648	1.840	-3.060	2.064	1.254	0.568	-8.458	0.615	1800
	(10.803***)	(23.146***)	(-2.896***)	(9.178***)	(6.486***)	(14.858***)	(-8.476***)		

Table 4: This table shows the results for the cross-sectional regressions, where proposed liquidity measure is regressed on bond characteristics and trading activity covariates. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit. The liquidity measure is based on the price dispersion and represents the root mean squared difference between TRACE prices and the respective Markit quotation calculated on a daily basis for each bond. TRACE prices are the actual traded prices and the Markit quotation represents the average market valuation, which we interpret as the market's expectation of the price. Thus, the proposed measure represents the volatility of the observed price deviations. The cross-sectional linear regressions use time-weighted averages of all variables. The results based on the whole time period, as well as based on each available quarter (2004 Q4 to 2006 Q3), are presented.

31

	Constant	Price Dispersion	R^2	Observations
2004 Q4	-18.192	0.021	0.220	1169
	(-220.88***)	(18.18***)		
2005 Q1	-18.796	0.028	0.239	1308
	(-207.59***)	(20.29***)		
2005 Q2	-18.344	0.022	0.234	1379
	(-228.37***)	(20.52***)		
2005 Q3	-18.296	0.024	0.244	1417
	(-248.61***)	(21.38***)		
2005 Q4	-17.838	0.018	0.198	1439
	(-261.99***)	(18.88***)		
2006 Q1	-18.436	0.026	0.229	1445
	(-243.79***)	(20.74***)		
2006 Q2	-18.487	0.031	0.313	1464
	(-259.91***)	(25.85***)		
2006 Q3	-18.377	0.027	0.273	1426
	(-262.90***)	(23.14***)		
Overall	-17.932	0.0251	0.313	1800
	(-321.51***)	(28.66***)		

Table 5: This table shows the results for the cross-sectional regressions, where the logarithmic Amihud measure is regressed on the proposed liquidity measure. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit. The liquidity measure is based on the price dispersion and represents the root mean squared difference between TRACE prices and the respective Markit quotation calculated on a daily basis for each bond. TRACE prices are the actual traded prices and the Markit quotation represents the average market valuation, which we interpret as the market's expectation of the price. Thus, the proposed measure represents the volatility of the observed price deviations. The cross-sectional linear regressions use time-weighted averages of all variables. The results based on the whole time period, as well as based on each available quarter (2004 Q4 to 2006 Q3), are presented.

	Constant	Price Dispersion	Coupon	Amount Issued	Trades	Trade Volume	R^2	Observations
2004 Q4	-10.033	0.015	0.190	-0.156	-0.049	-0.366	0.425	1169
	(-6.450***)	(12.835***)	(7.258***)	(-1.837*)	(-9.409***)	(-6.891***)		
2005 Q1	-1.567	0.016	0.223	-0.447	-0.041	-0.575	0.577	1308
	(-1.120)	(13.459***)	(9.707***)	(-5.959***)	(-12.445***)	(-12.276***)		
2005 Q2	-0.585	0.013	0.208	-0.568	-0.043	-0.454	0.602	1379
	(-0.479)	(14.990***)	(10.190***)	(-8.643***)	(-15.547***)	(-10.176***)		
2005 Q3	-8.183	0.016	0.168	-0.157	-0.098	-0.475	0.576	1417
	(-6.640***)	(16.563***)	(8.792***)	(-2.465**)	(-18.190***)	(-10.571***)		
2005 Q4	-5.783	0.014	0.156	-0.241	-0.045	-0.519	0.468	1439
	(-4.550***)	(15.310***)	(7.457***)	(-3.667***)	(-11.462***)	(-10.921***)		
2006 Q1	-7.904	0.020	0.186	-0.165	-0.093	-0.499	0.554	1445
	(-5.926***)	(17.550***)	(9.107***)	(-2.442**)	(-18.153***)	(-10.224***)		
2006 Q2	-8.202	0.024	0.140	-0.230	-0.109	-0.371	0.623	1464
	(-6.950***)	(22.752***)	(7.517***)	(-3.844***)	(-21.190***)	(-9.167***)		
2006 Q3	-8.464	0.021	0.198	-0.226	-0.059	-0.375	0.592	1426
	(-7.046***)	(20.535***)	(10.118***)	(-3.701***)	(-19.028***)	(-9.843***)		
Overall	-7.624	0.018	0.186	-0.275	-0.067	-0.336	0.637	1800
	(-8.600***)	(23.184***)	(13.391***)	(-5.759***)	(-20.976***)	(-8.764***)		
Overall w/o measure	0.177	-	0.296	-0.583	-0.050	-0.436	0.528	1800
	(0.189)	-	(19.909***)	(-11.128***)	(-14.065***)	(-10.027***)		

Table 6: This table shows the results for the cross-sectional regressions, where the logarithmic Amihud measure is regressed on the proposed liquidity measure, as well as on bond characteristics, and on trading activity variables. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit. The liquidity measure is based on the price dispersion and represents the root mean squared difference between TRACE prices and the respective Markit quotation calculated on a daily basis for each bond. TRACE prices are the actual traded prices and the Markit quotation represents the average market valuation, which we interpret as the market's expectation of the price. Thus, the proposed measure represents the volatility of the observed price deviations. The cross-sectional linear regressions use time-weighted averages of all variables. The results based on the whole time period, as well as based on each available quarter (2004 Q4 to 2006 Q3) are presented. The last row reports the results for the regression based on the whole time period, without using the proposed liquidity measure as explanatory variable, therefore, quantifying the reduction in R^2.
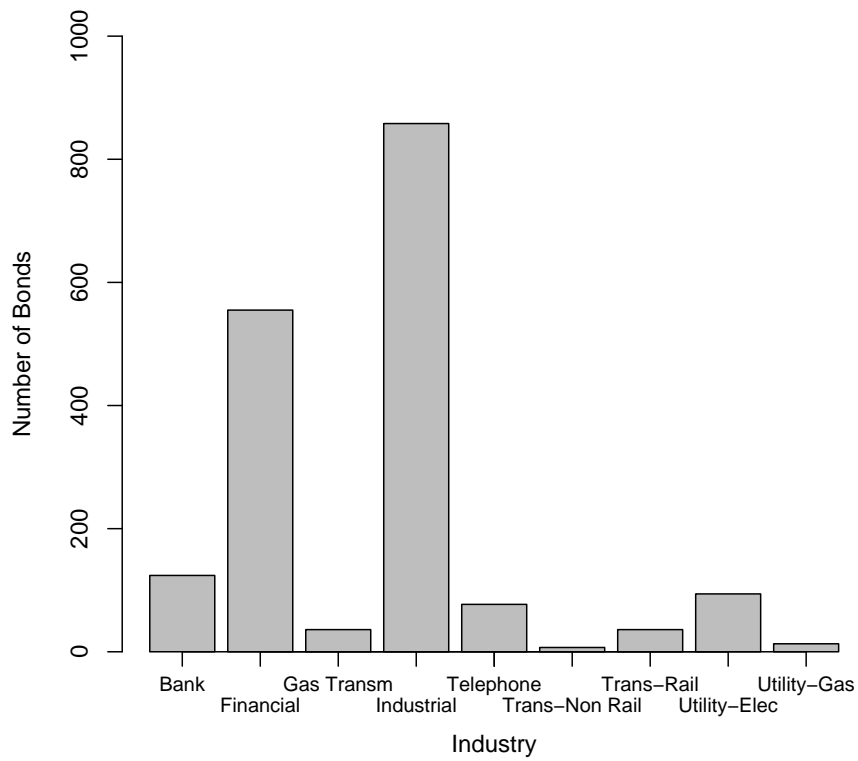
33

Figure 1: This figure shows the distribution of the corporate bonds in our data set across industries, which are represented by the Bloomberg issuer industry categories. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit.
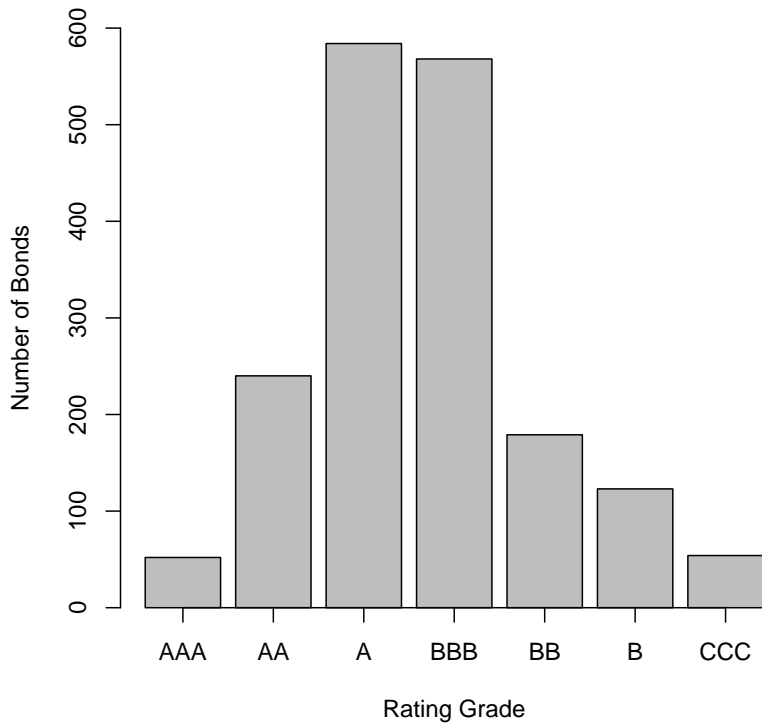
Figure 2: This figure shows the distribution of corporate bonds in our data set across rating grades. The grades represent the issue ratings and are taken from Fitch, Standard & Poor's, and Moody's. Each bond in the data set has at least one rating from one agency. If a bond has more than one rating, the grade represents the "average" rating. This average is generated by first transforming the individual ratings into numeric values (AAA=1 to CCC=7) and then using the rounded mean as average rating. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit.
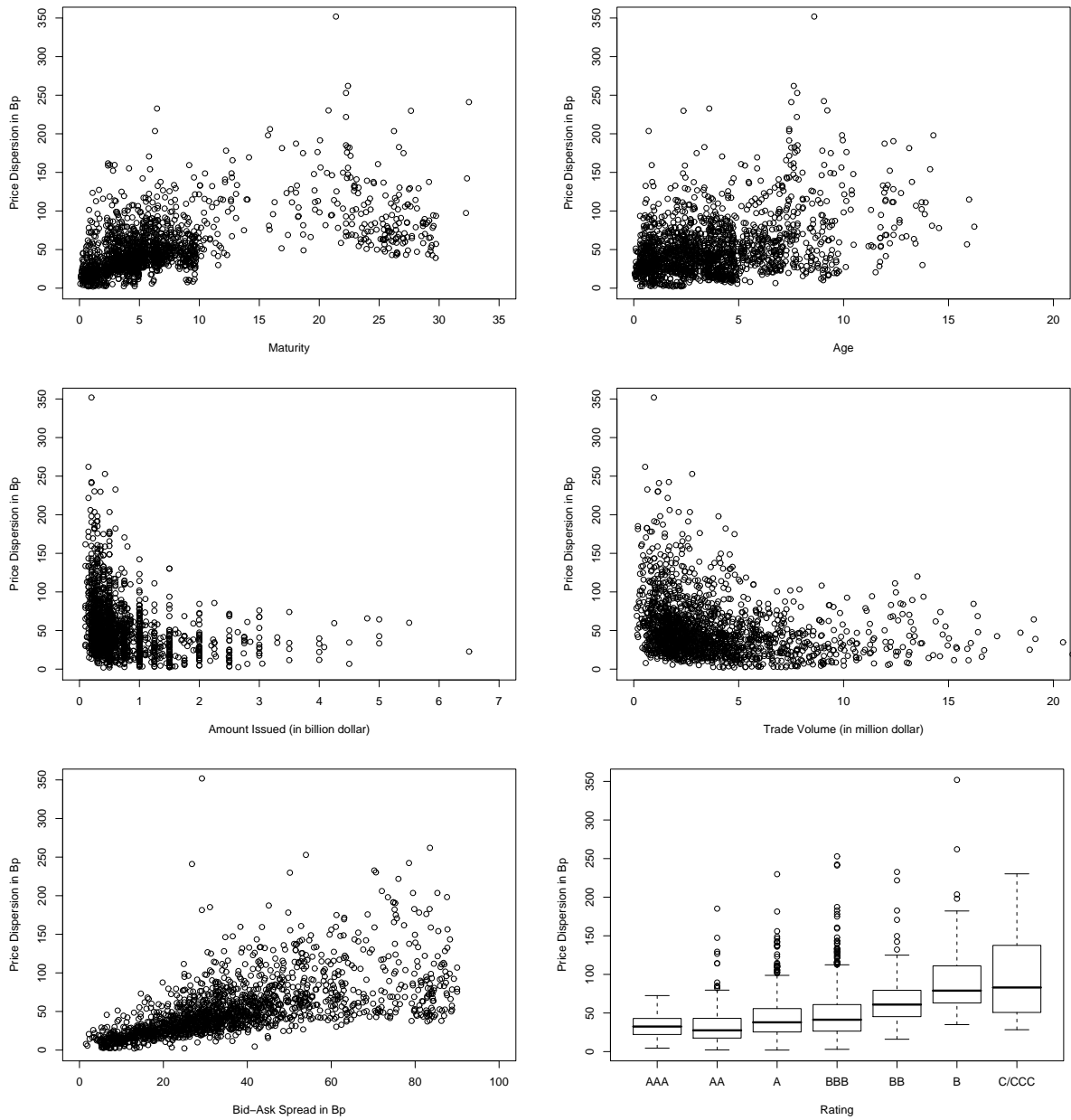
Figure 3: This figure analyzes the relation between the proposed liquidity measure and six bond characteristics and trading activity variables, i.e., maturity, age, amount issued, trade volume, bid-ask spread, and rating. The liquidity measure is based on the price dispersion and represents the root mean squared difference between TRACE prices and the respective Markit quotation calculated on a daily basis for each bond. TRACE prices are the actual traded prices and the Markit quotation represent the average market valuation, which we interpret as the market's expectation of the price. Thus, the proposed measure represents the volatility of the observed price deviations. In each graphical representation, one point corresponds to one bond represented by the time-series average value of its price dispersion measure and the relevant variable. As rating is an ordinal variable, boxplots for each rating grade are added to the graphical representation. The data set consists of a matched sample of 1,800 US corporate bonds with 3,889,017 transaction prices during the period October 2004 to October 2006, based on data from TRACE and Markit.